ON RECORD INDICES AND RECORD TIMES

Konstantin BOROVKOV[1)]

Steklov Mathematical Institute, Moscow, Russia and
Carl von Ossietzky Universität, Oldenburg, Germany

Dietmar PFEIFER

Carl von Ossietzky Universität, Oldenburg, Germany

Abstract: In this paper conditions are characterized under which the record indices of a sequence of independent, non–identically distributed random variables (with not necessarily continuous distribution functions) are independent, or, equivalently, the corresponding record times form a Markov chain with conditionally invariant transition probabilities (CITP chain). In particular, some former results of Nevzorov are simplified and generalized. The correlation structure of record indices in the general case is also studied.

AMS Subject Classifications: Primary 60G70, 62E10; Secondary 60J05

Key words and phrases: Record index, record time, Nevzorov scheme, maxima, CITP chains, partial correlation.

1. Introduction

Precisely 40 years ago, the very first paper – to our knowledge – on record values and record times appeared by Chandler (1952). Since then, a vast literature on the topic has emerged; see e.g. the latest survey paper by Nevzorov (1988). Among all of these the pioneering paper surely was the one by Rényi (1962) which established the surprising fact that the record indices of an i.i.d. sequence are independent (although this result is already contained in Dwass (1960), as is pointed out in Galambos (1987), p. 357, and Resnick (1987), p. 163). Generalizations to k–records are due to Ignatov's Theorem (cf. Resnick (1987), Chapter 4.6, or the review of Nagaraja (1988)). An elementary approach to the record index problem (for the i.i.d. case) is given in Pfanzagl (1991). To be more precise, let $X_1, \ldots, X_n$ be arbitrary real–valued random variables (r.v.'s), and denote $\overline{X}_k = \max_{j \leq k} X_j$. The random variables

$$I_j = \begin{cases} 1 & \text{if } X_j > \overline{X}_{j-1}, \\ 0 & \text{otherwise,} \end{cases} \quad j \geq 2, \qquad \text{and} \quad I_1 = 1,$$

are called *record indices* for the random vector $\mathbf{X} = (X_1, \ldots, X_n)$. The corresponding *record times* are the successive stopping times for $I_1, \ldots, I_n$, i.e. the instants τ_k at which $I_{\tau_k} = 1$. The probabilistic structure of record times in the case of infinite

[1)] Supported by the Alexander von Humboldt Foundation

i.i.d. sequences has been investigated by several authors, using different techniques, such as Deheuvels (1983), Galambos and Seneta (1975), Gut (1990), Pfeifer (1985, 1986), Pfeifer and Zhang (1989), among others. Vervaat's (1973) paper considers, in particular, the case of discrete random variables $\{X_k\}$. In most of these papers, the independence of record indices as stated by Rényi or, equivalently, the CITP structure (**c**onditionally **i**nvariant **t**ransition **p**robabilities) of the Markov chain of record times plays a central role. The same structure was also discovered in more complicated situations, first by Yang (1975), and later by Ballerini and Resnick (1985, 1987) and Nevzorov (see his 1988 survey paper). Since in the general case of independent r.v.'s the sequence of record times may become defective, it is surely more convenient to deal with the record indices instead, which will always exist as a proper sequence if $\boldsymbol{X}$ is infinite.

The independence of record indices also is an essential prerequisite for a suitable Poisson approximation of the number of records in n observations, i.e. $\sum_{j=1}^{n} I_j$. Applications of this approximation in connection with the so–called secretary problem and searching strategies in computer science are given, for example, in Pfeifer (1989, 1991).

Therefore, an essential problem is to characterize conditions under which record indices for independent, but otherwise arbitrary random variables are independent. However, in the general case, the character of dependence between record indices is very complicated and difficult to describe. The present paper deals with some properties of the joint distribution of the record indices as well.

2. Characterization of the case of independent record indices

The first deep result on the distribution of record indices is seemingly the above–mentioned theorem by Rényi (1962), which states that, if $X_1, \ldots, X_n$ are i.i.d.r.v.'s with continuous distribution function, then $I_1, \ldots, I_n$ are independent Bernoulli r.v.'s with success probabilities

$$p_k = \boldsymbol{P}(I_k = 1) = k^{-1}, \quad k = 1, \ldots, n.$$

On the other hand, in the case of non-identically independently distributed X_i's, the record indices need not to be independent (for more comments and references, see the survey paper by Nevzorov (1988)).

In Yang (1975), it was observed that, if we start with an i.i.d. sequence and form a new sequence, consisting of maxima of the original X_i's over non-overlapping segments of the original sequence, then the record indices for the new sequence will be independent too. Later, in a generalization of this scheme, Nevzorov proved the following result.

Theorem 1. *Let $X_1, \ldots, X_n$ be independent r.v.'s, $F_j(x) = \boldsymbol{P}(X_j < x)$ be their distribution functions, and suppose that*

$$F_j(x) = F^{\alpha_j}(x), \quad x \in \mathbf{R}, \ \alpha_j > 0, \ j = 1, \ldots, n, \tag{1}$$
where F is a fixed continuous distribution function.

Then $I_1, \ldots, I_n$ are independent r.v.'s with $p_k = \alpha_k \Big/ \sum_{j \le k} \alpha_j$.

We give a new proof of this result which uses, in fact, the following lemma describing to some extent general relations between probabilities p_k and β_k, where

$$\beta_k = \beta_{k,n} = \boldsymbol{P}(M_n = k)$$

is the probability of having the maximum of $X_1, \ldots, X_n$ on the kth position first, with

$$M_j = \min\{k \le j : \ X_k = \overline{X}_j\}$$

being the first index of the maximum of $X_1, \ldots, X_j$. Note that $\sum_{k \le n} \beta_k = 1$, since always $M_n \le n$.

Lemma 1. *Let $X_1, \ldots, X_n$ be arbitrary r.v.'s. If, for some $k \le n$, the events $\{I_k = 1\}$ and $\{I_{k+1} = I_{k+2} = \ldots = I_n = 0\}$ are independent, then, for this k,*

$$p_k := \boldsymbol{P}(I_k = 1) = \frac{\beta_k}{\sum_{j \le k} \beta_j}, \quad \text{if} \ \sum_{j \le k} \beta_j > 0. \tag{2}$$

Proof. Clearly, $\{I_j = 1\} = \{M_j = j\}$ for all $j = 1, \ldots, n$, and hence

$$\begin{aligned} \beta_k =& \boldsymbol{P}(M_n = k) = \boldsymbol{P}(I_k = 1, \ I_{k+1} = \ldots = I_n = 0) = \\ &= \boldsymbol{P}(I_k = 1) \ \boldsymbol{P}(I_{k+1} = \ldots = I_n = 0) = \boldsymbol{P}(I_k = 1) \ \boldsymbol{P}(M_n \le k) = p_k \sum_{j \le k} \beta_j, \end{aligned}$$

which proves Lemma 1. ∎

Note again that there are no assumptions on continuity of distribution functions in Lemma 1.

Proof of Theorem 1, in three steps:

(i) *Case $\alpha_j \equiv 1$.* This is just Rényi's (1962) famous result. Here independence of I_j follows immediately from the fact that order statistics and ranks are independent of each other in the i.i.d. case, and the values of p_k are obvious from Lemma 1 ($\beta_k = n^{-1} \equiv \alpha_k / n$ by symmetry).

(ii) *Case $\alpha_j = a_j/m$, a_j and m are integers, $j = 1, \ldots, n$.* Here without loss of generality we may assume that

$$X_1 = \max(Y_1, \ldots, Y_{a_1}), \ X_2 = \max(Y_{a_1+1}, \ldots, Y_{a_1+a_2}), \ldots,$$

where the Y_i are i.i.d. r.v.'s with common distribution function $F^{1/m}$ (cf. Yang (1975)). Clearly,

$$I_1 = \max(J_1, \ldots, J_{a_1}),\ I_2 = \max(J_{a_1+1}, \ldots, J_{a_1+a_2}), \ldots,$$

where the J_k are the record indices for $Y_1, \ldots, Y_N$ with $N = a_1 + \ldots + a_n$. Independence of the I_i is now obvious from that of the J_k (see (i)), with $\beta_k = a_k/N = \alpha_k m/N$.

(iii) *General case.* Here we just use the limit passage in (ii): let $a_{m,j} = \lfloor \alpha_j m \rfloor$ ($\lfloor \cdot \rfloor$ denoting integer part), then $\alpha_{m,j} := a_{m,j}/m \to \alpha_j$ as $m \to \infty$, and $F^{\alpha_{m,j}} \to F^{\alpha_j}$. Suppose that $X_{m,1}, \ldots, X_{m,n}$ are independent r.v.'s with distribution functions $F^{\alpha_{m,1}}, \ldots, F^{\alpha_{m,n}}$. Then the joint distribution of $(X_{m,1}, \ldots, X_{m,n})$ converges weakly to that of $(X_1, \ldots, X_n)$, and the assertion of Theorem 1 follows from (ii) and the continuity of F (which implies that the boundary of any of the sets $\{I_1 = \delta_1, \ldots, I_n = \delta_n\}$, $\delta_j = 0$ or 1, is a null–set with respect to the distribution of $(X_1, \ldots, X_n)$). ∎

Now, in view of Theorem 1, there arises a natural question, whether the independence of $I_1, \ldots, I_n$ implies (1). The answer is, in general, negative; however, the following partial characterization was given by Nevzorov (for details and references see his survey).

Theorem 2. *Let the r.v.'s $X_1, \ldots, X_n$ be independent and their distribution functions F_j have densities f_j with $\prod_{j=1}^n f_j(x) \neq 0,\ x \in (\alpha, \beta)$, for some $-\infty \le \alpha < \beta \le \infty$. If, for any r.v. X_{n+1}, which is independent of $X_1, \ldots, X_n$ and has an arbitrary density, I_{n+1} and $(I_1, \ldots, I_n)$ are independent, then* (1) *holds true.*

Remark 1. In fact, the proof of Theorem 2 implicitly makes use of the following property:

$$I_j \text{ and } \overline{X}_j \text{ are independent}, \quad j \le n, \tag{3}$$

which follows from the conditions of this theorem. Note also that under the assumptions of Theorem 1 property (3) is obvious. Indeed, for any x,

$$\begin{aligned} \boldsymbol{P}(\overline{X}_j < x,\ I_j = 1) &= \int_{-\infty}^{x} \left[\prod_{k<j} F_k(y) \right] dF_j(y) = \\ &= \frac{\alpha_j}{a(j)} F^{a(j)}(x) = p_j \boldsymbol{P}(\overline{X}_j < x), \quad a(j) = \sum_{k \le j} \alpha_j. \end{aligned} \tag{4}$$

Moreover, it is worth noting that the independence of $I_1, \ldots, I_n$ follows readily from this property, for $I_{j+1}, \ldots, I_n$ are functions of $(\overline{X}_j, X_{j+1}, \ldots, X_n)$ alone, and hence, from the independence of $X_1, \ldots, X_n$, we immediately see that I_j is independent of $(I_{j+1}, \ldots, I_n)$. Therefore, $I_1, \ldots, I_n$ are completely independent. Below we shall see that just this property (3) is characteristic for the case (1), when $X_1, \ldots, X_n$ are independent.

To formulate our result, we first introduce several notations. Let

$$D_k = \{x : \ \Delta F_k(x) > 0\}, \quad \Delta F(x) = F(x+0) - F(x),$$

be the set of all atoms of F_k, and let

$$W_j = \{x : \ H_j(x+0) > 0\} = \{x : \ \boldsymbol{P}(\overline{X}_{j-1} \le x) > 0\},$$

where $H_j(x) = \boldsymbol{P}(\overline{X}_{j-1} < x) = \prod_{k<j} F_k(x)$, $j > 1$. We shall use the following condition

$$[D_j] : \quad D_k \cap D_j \cap W_j = \emptyset \quad \text{for all} \ \ k < j.$$

Theorem 3. *Let $X_1, \ldots, X_n$ be independent r.v.'s. Then relation* (1) *is fulfilled if and only if condition $[D_j]$ holds for some $j > 1$ and, for all $j = 1, \ldots, n$, the r.v.'s I_j and $\overline{X}_j$ are independent with $0 < p_j < 1$.*

Remark 2. Note that no continuity of F_j is required. Condition $[D_j]$ is essential for (1), as is shown by the following example. Let

$$X_1 = \begin{cases} 1 & \text{with probability } 1-p, \\ 0 & \text{with probability } p, \end{cases} \qquad 0 < p < 1, \quad X_2 \equiv 1.$$

Then I_1, I_2 and $\overline{X}_2$ are independent, $p_2 = p$, but (1) does not hold. Note also that condition $[D_j]$ will obviously be satisfied, if F_j is continuous (in this case, $D_j = \emptyset$).

Now turn to the more general case, when one can also have $p_j = 0$ or 1. Denote

$$m_1 = 1, \quad m_{i+1} = \min\{k > m_i : \ p_k = 1\},$$

$$\Delta_i = \{j : \ m_i \le j < m_{i+1}\}, \quad \Delta_i^* = \{j \in \Delta_i : \ 0 < p_j < 1\}.$$

It is clear that $p_{m_i} = 1$ implies, together with independence of $\{X_j\}$, the independence of the record sequences on different Δ_i, so that the sequences $\{I_j, \ j \in \Delta_i\}$, $i = 1, 2, \ldots$, are independent, and it is easy to see that

$$\{x : \ H_{m_i} < 1\} \cap \{x : \ F_{m_i}(x+0) > 0\} = \emptyset.$$

Therefore we have from our Theorem 3 the following

Corollary 1. *Let $X_1, X_2, \ldots$ be independent r.v.'s. If, for some $i \ge 1$, I_j and $\overline{X}_j$ are independent of each other for all $j \in \Delta_i$ (or Δ_i^*), and $[D_j]$ holds for some $j \in \Delta_i^* \ne \emptyset$, then*

$$F_j(x) = F_{m_i}^{\alpha_j}(x), \quad j \in \Delta_i^*,$$

and F_{m_i} is continuous, whereas

$$F_j(x_{m_i} + 0) = 1, \quad j \in \Delta_i \backslash \Delta_i^*,$$

where $x_{m_i} = \inf\{x : \ F_{m_i}(x) > 0\}$. Moreover, $F_j(x_{m_i}) = 1, \ j < m_i$.

Proof of Theorem 3. By Theorem 1 and Remark 1, it remains only to prove that if $[D_j]$ holds for some $j > 1$, and I_j and $\overline{X}_j$ are independent for all $j > 1$, then (1) is fulfilled.

Denote $y_j = \inf W_j$, $x_j = \inf\{x : \ F_j(x) > 0\}$. We show now that $y_j = x_j$. Similary to (4), we have, by the independence of I_j and $\overline{X}_j$, that

$$L_j(x) := \int_{-\infty}^{x} H_j(y)\, dF_j(y) = p_j H_j(x) F_j(x) =: R_j(x), \quad x \in \mathbf{R}. \tag{5}$$

Suppose now that $y_j < x_j$ which means that $h = H_j(x_j+0) > 0$. Now if $\Delta F_j(x_j) = 0$, then

$$L_j(x) \sim hF_j(x), \quad R_j(x) \sim p_j h F_j(x) \quad \text{as} \quad x \downarrow x_j. \tag{6}$$

If $\Delta F_j(x_j) > 0$, then by condition $[D_j]$ we have $\Delta F_k(x_j) = 0$ for all $k < j$ and therefore $\Delta H_j(x_j) = 0$, hence once again we come to (6). But (6) is clearly a contradiction to (5), for $0 < p_j < 1$ and $F_j(x) > 0$, $x > x_j$. Hence $y_j \geq x_j$.

Now let $y_j > x_j$ which means that $f = F_j(y_j) > 0$. If $h := H_j(y_j + 0) > 0$, then by condition $[D_j]$ we have $\Delta F_j(y_j) = 0$, and hence

$$L_j(x) \to 0, \quad R_j(x) \geq p_j h f > 0 \quad \text{as } x \downarrow y_j, \tag{7}$$

since $p_j > 0$. If $h = 0$, then

$$L_j(x)/H_j(x) \to 0, \quad R_j(x) \geq p_j H_j(x) f \quad \text{as } x \downarrow y_j, \tag{8}$$

where the first relation follows from the fact that here

$$L_j(x) = \int_{y_j+0}^{x} H_j(y) dF_j(y) \leq H_j(x)\big(F_j(x) - F_j(y_j + 0)\big).$$

Thus in both cases (7) and (8) we have a contradiction to (5), and hence finally $y_j = x_j =: z$.

Since

$$\Delta H_j(x) \cdot \Delta F_j(x) \equiv 0 \tag{9}$$

by condition $[D_j]$, we have the following relations for the measures involved:

$$H_j(z + 0)\, F_j(z + 0) = 0, \tag{10}$$

$$d(H_j\, F_j) = H_j\, dF_j + F_j\, dH_j, \tag{11}$$

and hence it follows from (5) that

$$(1 - p_j) H_j(x) dF_j(x) = p_j F_j(x) dH_j(x), \quad x > z,$$

and therefore

$$\frac{dF_j(x)}{F_j(x)} = \gamma_j \frac{dH_j(x)}{H_j(x)}, \quad x > z, \quad \gamma_j = \frac{p_j}{1 - p_j}. \tag{12}$$

Hence the measures with distribution functions $F_j(x)$ and $H_j(x)$ are equivalent on $\{x > z\}$, and therefore, by virtue of (9), are both continuous there. From this and (12) we obtain

$$d \log F_j(x) = \gamma_j d \log H_j(x), \quad x > z,$$

and therefore that

$$F_j(x) = H_j^{\gamma_j}(x) + C, \quad x > z.$$

But obviously $C = 0$ (let $x \to \infty$), and from (10) it follows that $F_j(z+0) = H_j(z+0) = 0$ and hence

$$F_j(x) = H_j^{\gamma_j}(x) \quad \text{for all} \quad x \in \mathbf{R}.$$

Further, since $H_j(x) > 0, \ x > z$, and is continuous, we have that, for all $k < j$, $F_k(x)$ and $H_k(x)$ are positive and continuous on $\{x > z\}$, too. Hence (11) and (12) also hold for all $k \le j$, so that

$$F_k(x) = H_k^{\gamma_k}(x), \quad x > z, \quad k \le j,$$

and therefore

$$F_k(x) = F_1^{\alpha_k}(x), \quad x > z, \quad k \le j,$$

$\alpha_k = p_k \Big/ \prod_{m=2}^{k} (1 - p_m)$. Since $F_j(z+0) = 0$, this representation holds for all $x \in \mathbf{R}$.

In case that $j = n$, the theorem is already proved. Otherwise we shall show how to derive the desired representation for all $k > j$.

Since $H_{j+1}(x) = H_j(x)\, F_j(x) > 0$ for $x > z$ and is continuous, we obtain easily from (5) (with $j+1$ instead of j) that

$$p_{j+1} F_j^{1/\gamma_j + 1}(x) F_{j+1}(x) = \int_{-\infty}^{x} F_j^{1/\gamma_j + 1}(y) dF_{j+1}(y)$$

and therefore

$$\frac{dF_{j+1}}{F_{j+1}} = \frac{p_{j+1}}{p_j(1 - p_{j+1})} \frac{dF_j}{F_j}, \quad x > \max(z, x_{j+1}).$$

This means that

$$F_{j+1}(x) = F_1^{\alpha_{j+1}}(x), \quad x > \max(z, x_{j+1}).$$

Clearly, the case of $x_{j+1} > z$ and $F_{j+1}(x_{j+1}) = 0$ is impossible. Further, if $x_{j+1} \le z$, we immediately obtain the required representation for all $x \in \mathbf{R}$. Now there remains only the case $x_{j+1} > z, \ f = F_{j+1}(x_{j+1} + 0) > 0$. But in this situation

$$L_{j+1}(x) \to hf, \quad R_{j+1}(x) \to p_{j+1} hf \quad \text{as} \ \ x \to x_{j+1} + 0,$$

where $h = H_{j+1}(x_{j+1}) > 0$, since $x_{j+1} > z$. So we obtain again a contradiction to (5) (with $j+1$ instead of j). Hence $F_{j+1}(x) = F_1^{\alpha_{j+1}}(x)$ for all $x \in \mathbf{R}$, and the same argument applies for $j+2$ etc. Thus Theorem 3 is proved. ■

3. Record times and the CITP property

In this section we shall establish a relationship between independent record indices and record times with the CITP property. For this purpose, consider a (possibly defective) homogeneous Markov chain $\{T_m\}_{m=1}^{\infty}$ with $T_1 = 1$ and values in $\mathbf{N} \cup \{\infty\}$ which is strictly increasing on finite values and has ∞ as absorbing state. We say that $\{T_m\}$ possesses the CITP(n) property with $n \in \mathbf{N}$ (conditionally invariant transition probability), if for all $j \le n$ the ratio

$$p_{n+1} := \frac{\boldsymbol{P}(T_{m+1} = n+1 \mid T_m = j)}{\boldsymbol{P}(T_{m+1} > n \mid T_m = j)} \tag{13}$$

is independent of j whenever the denominator is positive. Note that under the above conditions, there always exists some j such that the denominator is positive, e.g. for $j = n$ in which case $\boldsymbol{P}(T_{m+1} > n \mid T_m = j) = 1$, such that p_{n+1} is well–defined (possibly with value zero). If $\{T_m\}$ is CITP(n) for all $n \in \mathbf{N}$, then $\{T_m\}$ is called a CITP chain. Note that the validity of (13) also implies

$$p_{n+1} = \boldsymbol{P}(T_{m+1} = n+1 \mid T_m \le n < T_{m+1}) \tag{14}$$

since (13) is equivalent to

$$\boldsymbol{P}(T_{m+1} = n+1 \mid T_m = j) = p_{n+1}\boldsymbol{P}(T_{m+1} > n \mid T_m = j) \tag{15}$$

for all $j \le n$, and hence also

$$\boldsymbol{P}(T_{m+1} = n+1, T_m = j) = p_{n+1}\boldsymbol{P}(T_{m+1} > n, T_m = j)$$

for all $j \le n$ from which (14) follows by summation.

For a homogeneous Markov chain $\{T_m\}$ as above (not necessarily CITP) let the sequence $\{J_n\}$ given by

$$J_n = \begin{cases} 1, & \text{if } T_m = n \text{ for some } m \\ 0, & \text{otherwise} \end{cases} \qquad (n \in \mathbf{N}) \tag{16}$$

denote the associated occurrence indices. The following result provides a connection between such chains and partial independence of the occurrence indices.

Lemma 2. *$\{T_m\}$ is CITP(n) for some $n \in \mathbf{N}$ if and only if J_{n+1} and $(J_1, \dots, J_n)$ are independent. In this case,*

$$\boldsymbol{P}(J_{n+1} = 1) = \frac{\boldsymbol{P}(T_{m+1} = n+1 \mid T_m = j)}{\boldsymbol{P}(T_{m+1} > n \mid T_m = j)} = \boldsymbol{P}(T_{m+1} = n+1 \mid T_m \le n < T_{m+1})$$

for all $j \le n$ such that the denominator is positive.

Proof. Let $i_1, \dots, i_n \in \{0,1\}$ with $i_1 = 1$ and $m = \sum_{\ell=1}^{n} i_\ell$. Let further $j_1 < j_2 < \dots < j_m$ denote those indices w.r.t. $(i_1, \dots, i_n)$ such that $i_{j_\ell} = 1$. (In particular, $j_1 = 1$.) Then, by homogeneity, we have

$$\begin{aligned} \boldsymbol{P}(J_{n+1} = 1, J_n = i_n, \dots, J_1 = i_1) &= \boldsymbol{P}(T_{m+1} = n+1, T_m = j_m, \dots, T_1 = j_1) \\ \boldsymbol{P}(J_n = i_n, \dots, J_1 = i_1) &= \boldsymbol{P}(T_{m+1} > n, T_m = j_m, \dots, T_1 = j_1). \end{aligned} \tag{17}$$

Suppose now that $\{T_m\}$ is CITP(n). Then by (15) and the Markov property, we have

$$\begin{aligned}&\boldsymbol{P}(T_{m+1}=n+1, T_m=j_m, \dots, T_1=j_1) = \\ &= p_{n+1}\boldsymbol{P}(T_{m+1}>n, T_m=j_m, \dots, T_1=j_1)\end{aligned} \tag{18}$$

and hence by (17),

$$\boldsymbol{P}(J_{n+1}=1, J_n=i_n, \dots, J_1=i_1) = p_{n+1}\boldsymbol{P}(J_n=i_n, \dots, J_1=i_1), \tag{19}$$

which means that J_{n+1} and $(J_1, \dots, J_n)$ are independent. Conversely, if J_{n+1} and $(J_1, \dots, J_n)$ are independent, then (19) holds for all choices of $i_1, \dots, i_n$ and hence also (18) for all choices of $j_1, \dots, j_m$. Backwards calculation then shows that (15) is valid; hence the lemma is proved. ∎

Note that in the case of i.i.d. r.v.'s $X_1, X_2, \dots$, the sequence $\{\tau_m\}$ of record times is a (homogeneous) CITP chain since we have

$$\boldsymbol{P}(\tau_{m+1}=k \mid \tau_m=j) = \frac{j}{k(k-1)}, \quad 1 \le j < k$$

(see e.g. Rényi (1962)), hence for $j \le n$,

$$p_{n+1} = \frac{\boldsymbol{P}(T_{m+1}=n+1 \mid T_m=j)}{\boldsymbol{P}(T_{m+1}>n \mid T_m=j)} = \frac{\dfrac{j}{(n+1)n}}{\dfrac{j}{n}} = \frac{1}{n+1}$$

as expected. For the scheme (1), we similarly have

$$\boldsymbol{P}(\tau_{m+1}=k \mid \tau_m=j) = \frac{\alpha_k \sum_{i=1}^{j}\alpha_i}{\left(\sum_{i=1}^{k-1}\alpha_i\right)\left(\sum_{i=1}^{k}\alpha_i\right)}, \quad 1 \le j < k,$$

hence $\{\tau_m\}$ here also is a CITP chain (see e.g. Pfeifer (1989)) with

$$p_{n+1} = \alpha_{n+1} \Big/ \sum_{i=1}^{n+1}\alpha_i.$$

By Theorems 2 and 3, the latter case is essentially the only one in which, for independent r.v.'s $\{X_m\}$, the record times are a CITP or CITP(n) chain.

Corollary 2. *Let the r.v.'s $\{X_m\}$ be independent and their first n distribution functions F_j have densities f_j with $\prod_{j=1}^{n} f_j(x) \neq 0$, $x \in (\alpha, \beta)$, for some $-\infty \le \alpha < \beta \le \infty$. If the record times $\{\tau_m\}$ are CITP(n), then (1) holds true.*

Proof. Immediate from Lemma 2 and Theorem 2. ∎

Remark 3. By Theorem 3, a necessary condition for (1) to hold true is $0 < p_j < 1$ for $j = 1, 2, \ldots, n$. In this case, a CIPT chain can be characterized by the following structure of transition probabilities:

$$\boldsymbol{P}(T_{m+1} > k \mid T_m = j) = \frac{Q(k)}{Q(j)}, \quad 1 \le j \le k,$$

with some non–decreasing function $Q:\ \mathbf{N} \to [0,1]$; in fact, $Q(1) = 1$ and for $k \ge 2$, $Q(k) = \prod_{\ell=2}^{k}(1 - p_\ell)$ with the p_ℓ being the success probabilities of the occurrence indicators. Note, however, that such a representation fails to hold in the degenerate case, i.e. $p_j \in \{0, 1\}$ for some $j \ge 2$. For example, in the i.i.d. case, $Q(k) = 1/k$, while in the general Nevzorov scheme (1),

$$Q(k) = 1 \Big/ \sum_{\ell=1}^{k} \alpha_\ell, \quad k \ge 2.$$

4. Bounds for correlations between record indices

In this section we suppose that $\boldsymbol{X} = (X_1, \ldots, X_n)$ is a vector of independent r.v.'s X_j and denote by

$$\boldsymbol{p} = \boldsymbol{p}(\boldsymbol{X}) = (p_1, \ldots, p_n), \quad p_j = p_j(\boldsymbol{X}) = \boldsymbol{P}(I_j = 1),$$

the vector of record probabilities for the sample $\boldsymbol{X}$ (recall that $p_1 = 1$). It is well known that, for an arbitrary non-negative vector $\boldsymbol{p}$ of $p_j \le 1,\ p_1 = 1$, there always exists a random vector $\boldsymbol{X}$ with independent components such that $\boldsymbol{p}(\boldsymbol{X}) = \boldsymbol{p}$. Such an example is provided e.g. by independent r.v.'s X_j following laws $F^{\alpha_j}(x)$, where $F(x)$ is a fixed continuous distribution function, and

$$\alpha_j = p_j/(1 - p_2)(1 - p_3) \cdot \ldots \cdot (1 - p_j), \quad j = 1, \ldots, n, \tag{20}$$

in the non-degenerate case of $\max_{j>1} p_j < 1$ (see Theorem 1); otherwise the construction of $\boldsymbol{X}$ consists of several such "blocks" separated by the points j with $p_j = 1$ (see Corollary 1 above). In this example $I_j,\ j = 1, \ldots, n$, are independent, so that

$$p_{i,j} = p_{i,j}(\boldsymbol{X}) \equiv \boldsymbol{P}(I_i = I_j = 1) = p_i p_j.$$

However, in the general case the I_j need not be independent. Therefore a natural question arises: what limitations on the joint distribution of $\boldsymbol{I} = (I_1, \ldots, I_n)$ are imposed by the fact that $\boldsymbol{I}$ is a vector of record indices for some sample $\boldsymbol{X}$ with independent components? Clearly, the set of such limitations is not empty (say, $p_j = 1$ implies that $(I_1, \ldots, I_{j-1})$ and $(I_{j+1}, \ldots, I_n)$ are independent), but no answer is known so far to this question.

Here we deal only with two-dimensional marginal distributions of $\boldsymbol{I}$, i.e. with probabilities $p_{i,j}$. We shall show below that there are no restrictions on the values of $p_{j,j+1}$, and of what kind could be restrictions on $p_{j,j+2}$.

Denote

$$b^-(x,\ y) = \max(0,\ x + y - 1), \quad b^+(x,\ y) = \min(x,\ y).$$

From the probability addition theorem ($\boldsymbol{P}(A \cup B) = \boldsymbol{P}(A) + \boldsymbol{P}(B) - \boldsymbol{P}(AB)$), one can easily see that

$$b^-(\boldsymbol{P}(A),\ \boldsymbol{P}(B)) \le \boldsymbol{P}(AB) \le b^+(\boldsymbol{P}(A),\ \boldsymbol{P}(B))$$

are sharp (i.e. attainable) bounds for the joint probability $\boldsymbol{P}(AB)$ of any two events A and B with fixed probabilities $\boldsymbol{P}(A)$ and $\boldsymbol{P}(B)$. Now we introduce also the functions

$$c^-(x,\ y;\ z) = \begin{cases} b^-(x,\ y), & xy \le v = (1-x)(1-z), \\ x(xy - v)/(x - v), & xy > v; \end{cases}$$

$$c^+(x,\ y;\ z) = \begin{cases} b^+(x,\ y), & z \le \max(x,\ y), \\ xy/z, & z > \max(x,\ y). \end{cases}$$

Denote, for a given $\boldsymbol{p}$ $(p_1 = 1,\ 0 \le p_j \le 1)$, by

$$p^-_{i,j}(\boldsymbol{p}) = \min_{\boldsymbol{X}} \{p_{i,j} :\ \boldsymbol{p}(\boldsymbol{X}) = \boldsymbol{p}\}, \quad p^+_{i,j}(\boldsymbol{p}) = \max_{\boldsymbol{X}} \{p_{i,j} :\ \boldsymbol{p}(\boldsymbol{X}) = \boldsymbol{p}\}$$

the minimal and the maximal probabilities, resp., of joint records on the ith and jth places for a fixed vector $\boldsymbol{p}$ of (marginal) probabilities.

Theorem 4. *For any given* $\boldsymbol{p}$ *with* $p_1 = 1,\ 0 \le p_j \le 1$,

i) $p^\pm_{i,i+1}(\boldsymbol{p}) = b^\pm(p_i,\ p_j)$;
ii) $p^-_{i-1,i+1}(\boldsymbol{p}) \le c^-(p_{i-1},\ p_{i+1};\ p_i)$, $\quad p^+_{i-1,i+1}(\boldsymbol{p}) \ge c^+(p_{i-1},\ p_{i+1};\ p_i)$.

Remark 4. The bounds for $p^\pm_{i-1,i+1}(\boldsymbol{p})$ are seemingly not final. However, at least in the case $p_{i-1} \le p_{i+1}$, the "critical value" $m := \max\ (p_{i-1},\ p_{i+1})$ in the definition of c^+ above is also the "critical point" for $p^+_{i-1,i+1}(\boldsymbol{p})$, i.e. the value $b^+(p_{i-1},\ p_{i+1})$ is no longer attainable, if $p_i > m$. Indeed, if

$$p_{i-1,i+1} = b^+(p_{i-1},\ p_{i+1}) = p_{i-1} \le p_{i+1}, \tag{21}$$

then

$$\{X_{i-1} > \max_{j<i-1} X_j\} \subseteq \{X_{i+1} > \max_{j<i+1} X_j\} \subseteq \{X_{i+1} > X_i\},$$

and, by the independence of X_j's, this implies that $X_{i+1} > X_i$ a.s., and therefore that $\{I_i = 1\} \subseteq \{I_{i+1} = 1\}$ and $p_i \le p_{i+1}$, and hence (21) cannot be true for $p_i > m$.

Proof of Theorem 4. i). First we shall give constructions, proving this part of the theorem for $n = 3,\ i = 2$. In i$^-$), one has $p_{2,3} = b^-(p_2,\ p_3)$, and in i$^+$), $p_{2,3} = b^+(p_2,\ p_3)$.

i$^-$). Let $X_1 \equiv 1$, $\boldsymbol{P}(X_2 = 0) = 1 - p_2$, $\boldsymbol{P}(X_2 = 4) = p_2$, $\boldsymbol{P}(X_3 = j) = a_j$, $j = 0, 1, \dots$.

Case $p_2 + p_3 \le 1$. Put $a_0 = 1 - a_2$, $a_2 = p_3/(1 - p_2)$. We have $\boldsymbol{P}(I_2 = 1) = p_2$ and

$$\boldsymbol{P}(I_3 = 1) = \boldsymbol{P}(X_2 = 0,\ X_3 = 2) = p_3, \quad \boldsymbol{P}(I_2 = I_3 = 1) = 0 = b^-(p_2,\ p_3).$$

Case $p_2 + p_3 > 1$. Put $a_2 = 1 - a_5$, $a_5 = (p_2 + p_3 - 1)/p_2$. Clearly, we now have

$$\boldsymbol{P}(I_3 = 1) = \boldsymbol{P}(X_3 = 5) + \boldsymbol{P}(X_2 = 0,\ X_3 = 2) = p_3,$$

$$\boldsymbol{P}(I_2 = I_3 = 1) = \boldsymbol{P}(X_2 = 4,\ X_3 = 5) = p_2 a_5 = p_2 + p_3 - 1 = b^-(p_2,\ p_3).$$

i$^+$). Let $\boldsymbol{P}(X_1 = 1) = 1 - \boldsymbol{P}(X_1 = 5) = a_1$, $\boldsymbol{P}(X_2 = j) = b_j$, $j = 0, 1, \dots;$ $X_3 \equiv 3$.

Case $p_2 \le p_3$. Put $a_1 = p_3$, $b_2 = 1 - b_0 = p_2/p_3$. Here,

$$\boldsymbol{P}(I_2 = 1) = \boldsymbol{P}(X_1 = 1,\, X_2 = 2) = p_3 b_2 = p_2, \quad \boldsymbol{P}(I_3 = 1) = \boldsymbol{P}(X_1 = 1) = p_3,$$

and $\{I_2 = 1\} = \{I_2 = I_3 = 1\}$, so that $\boldsymbol{P}(I_2 = I_3 = 1) = p_2 = b^+(p_2,\, p_3)$.

Case $p_2 > p_3$. Put $a_1 = p_3/b_2$, $b_2 = 1 - b_6 = 1 - p_2 + p_3$. Clearly,

$$\boldsymbol{P}(I_2 = 1) = \boldsymbol{P}(X_2 = 6) + \boldsymbol{P}(X_1 = 1,\, X_2 = 2) = b_6 + a_1 b_2 = p_2,$$

$$\boldsymbol{P}(I_3 = 1) = \boldsymbol{P}(X_1 = 1,\, X_2 = 2) = a_1 b_2 = p_3,$$

and $\{I_3 = 1\} = \{I_2 = I_3 = 1\}$, so that $\boldsymbol{P}(I_2 = I_3 = 1) = p_2 = b^+(p_2,\, p_3)$.

Now let us consider the case of $n > 3$, $i < n$. It is not hard to see how these constructions can be generalized to the general case. In i$^-$), the left part of the sample $\boldsymbol{X}$ (i.e. $X_1, \dots, X_{i-1}$) could be constructed in an arbitrary way with $X_j \in (1/2,\ 3/2)$ a.s. and with $\boldsymbol{P}(I_j = 1) = p_j$ (say, using Theorem 1 with continuous distribution function F having support in $(1/2,\ 3/2)$, see (20) and comments there). The right part of $\boldsymbol{X}$ can be easily added step by step, using the basic relations

$$p_j = \int H_j(x) dF_j(x), \quad H_j(x) = \prod_{k<j} F_k.$$

In i$^+$), it suffices to construct the left part of $\boldsymbol{X}$ so that

$$H_i(5.5) - H_i(4.5) = a_5, \quad H_i(1.5) - H_i(0.5) = a_1.$$

This can be done using once again (1) in the following way. Take as a basic distribution F the uniform one on $(0, 1)$. Then we get independent r.v.'s $Y_1, \dots, Y_{i-1}$ with

$$G_i(x) = \prod_{k<i} x^{\alpha_k} = x^{\alpha_1 + \dots \alpha_{i-1}}$$

being the distribution function of $\max_{k<i} Y_k$. It remains to put $X_k = f(Y_k)$, $k = 1, \dots, i-1$, where $f(x) = 0.5 + x + 4I[x > G^{-1}(a_1)]$. The right part of $\boldsymbol{X}$ can be constructed in the same way as in i$^-$).

ii). Here we restrict ourselves to constructions for $n = 4$, $i = 2$, for which ii$^-$) $p_{2,4} = c^-(p_2, p_4;\ p_3)$, and ii$^+$) $p_{2,4} = c^+(p_2, p_4;\ p_3)$. The general case can be treated exactly in the same way as in i).

ii$^-$). We shall take the constructions from the corresponding parts of i$^-$) as the basic ones and then add the fourth component X_4 to meet our requirements.

Case $p_2 + p_3 \le 1$. Here we add X_4 with the following distribution.

a) If $p_2 + p_4 \le 1$, then we put

$$\boldsymbol{P}(X_4 = 3) = 1 - \boldsymbol{P}(X_4 = 0) = p_4/(1 - p_2).$$

Then $\{I_4 = 1\} = \{X_2 = 0,\ X_4 = 3\}$, and hence $\boldsymbol{P}(I_4 = 1) = p_4$. Clearly, $\boldsymbol{P}(I_2 = I_4 = 1) = 0 = b^-(p_2, p_4)$.

b) If $p_2 + p_4 > 1$, then we put

$$\boldsymbol{P}(X_4 = 3) = 1 - \boldsymbol{P}(X_4 = 6) = (1 - p_4)/p_2.$$

Then $\boldsymbol{P}(I_4 = 1) = \boldsymbol{P}(X_4 = 6) + \boldsymbol{P}(X_2 = 0,\ X_4 = 3) = p_4$, and

$$\boldsymbol{P}(I_2 = I_4 = 1) = \boldsymbol{P}(X_2 = 3,\ X_4 = 6) = b^-(p_2, p_4).$$

Case $p_2 + p_3 > 1$. Let $v = (1 - p_2)(1 - p_3)$.

a) If $p_2 p_4 \le v$, we put

$$\boldsymbol{P}(X_4 = 3) = 1 - \boldsymbol{P}(X_4 = 0) = p_2 p_4 / v.$$

Then clearly $\boldsymbol{P}(I_2 = I_4 = 1) = 0 = b^-(p_2, p_4)$, the last equality follows from the relations $p_2 p_4 \le v < (1 - p_2)p_2$.

b) If $p_2 p_4 > v$, we put

$$\boldsymbol{P}(X_4 = 3) = 1 - \boldsymbol{P}(X_4 = 6) = (1 - p_4)/(1 - v/p_2).$$

Then $\boldsymbol{P}(I_2 = I_4 = 1) = p_2 \boldsymbol{P}(X_4 = 6) = c^-(p_2, p_4;\ p_3)$.

ii$^+$). *Case* $p_3 \le p_4$.

a) If $p_2 \le p_4$, we take the construction from i$^+$) with $a_1 = 1 - a_5 = p_4$, $b_2 = 1 - b_0 = p_2/p_4$, but with $X_4 \equiv 4$ and with $\boldsymbol{P}(X_3 = 3) = 1 - \boldsymbol{P}(X_3 = 0) = p_3/p_4$. Then $\boldsymbol{P}(I_4 = 1) = \boldsymbol{P}(X_1 = 1) = p_4$, $\boldsymbol{P}(I_3 = 1) = \boldsymbol{P}(X_1 = 1, X_3 = 3) = p_3$, and $\boldsymbol{P}(I_2 = I_4 = 1) = \boldsymbol{P}(I_2 = 1) = p_2 = b^+(p_2, p_4)$.

b) If $p_2 > p_4$, we proceed in a similar way, using the scheme of i$^+$), case $p_2 > p_3$.

Case $p_3 > p_4$.

a) If $p_2 \le p_3$, we take the construction from i$^+$), case $p_2 \le p_3$, and X_4 with $\boldsymbol{P}(X_4 = 4) = 1 - \boldsymbol{P}(X_4 = 0) = p_4/p_3$. Then we have

$$\boldsymbol{P}(I_4 = 1) = \boldsymbol{P}(I_3 = 1,\ X_4 = 4) = p_4,$$

and

$$\boldsymbol{P}(I_2 = I_4 = 1) = \boldsymbol{P}(X_1 = 1)\,\boldsymbol{P}(X_2 = 2)\,\boldsymbol{P}(X_4 = 4) = \\ = p_2 p_4/p_3 = c^+(p_2,\, p_4;\ p_3).$$

b) If $p_2 \ge p_3$, we use the same X_4, but with the scheme of i$^+$), case $p_2 > p_3$. Now

$$\boldsymbol{P}(I_2 = I_4 = 1) = \boldsymbol{P}(X_1 = 1)\,\boldsymbol{P}(X_2 = 2)\,\boldsymbol{P}(X_4 = 4) = p_4,$$

and this completes the proof of Theorem 4. ■

Acknowledgements. This joint work was enabled by a grant of the Alexander von Humboldt Foundation while the first author was visiting the Carl von Ossietzky Universität Oldenburg, which is gratefully appreciated.

We would also like to thank the referees for their detailed criticism which lead to a considerable improvement in the presentation of the paper.

References

Chandler, K.N. (1952). The distribution and frequency of record values. *J. Roy. Stat. Soc. Ser. B* **14**, 220 – 228.

Deheuvels, P. (1983). The complete characterization of the upper and lower class of the record and inter–record times of an i.i.d. sequence. *Z. Wahrsch. verw. Geb.* **62**, 1 – 6.

Dwass, M. (1960). Some k–sample rank order tests. *Contributions to probability and statistics.* Stanford Univ. Press, Stanford, Cal., 198 – 202.

Ballerini, R., Resnick, S.I. (1985). Records from improving populations. *J. Appl. Prob.* **22**, 487 – 502.

Ballerini, R., Resnick, S.I. (1987). Records in the presence of a linear trend. *Adv. Appl. Prob.* **19**, 801 – 828.

Galambos, J. (1987). *The Asymptotic Theory of Extreme Order Statistics.* 2nd ed., Krieger Publ. Comp., Malabar, Florida.

Galambos, J., Seneta, E. (1975). Record times. *Proc. Amer. Math. Soc.* **50**, 383 – 387.

Gut, A. (1990). Limit theorems for record times. *Probab. Th. Math. Stat.* **1**, 490 – 503.

Gut, A. (1990). Convergence rates for for record times and the associated counting process. *Stoch. Proc. Appl.* **36**, 135 – 151.

Nagaraja, H.N. (1988). Record values and related statistics, a review. *Comm. Stat. Th. Meth.* **17**, 2223 – 2238.

Nevzorov, V.B. (1988). Records. *Th. Probab. Appl.* **32**, 201 – 228.

Pfanzagl, J. (1991). *Elementare Wahrscheinlichkeitsrechnung.* 2nd ed., W. de Gruyter, Berlin.

Pfeifer, D. (1985). On a joint strong approximation theorem for record and inter–record times. *Th. Probab. Rel. Fields* **75**, 213 – 221.

Pfeifer, D. (1986). Extremal processes, record times and strong approximation. *Publ. Inst. Stat. Univ. Paris* **XXXI**, 47 – 65.

Pfeifer, D. (1989). Extremal processes, secretary problems and the 1/e law. *J. Appl. Prob.* **26**, 722 – 733.

Pfeifer, D. (1991). Some remarks on Nevzorov's record model. *Adv. Appl. Prob.* **23**, 823—834.

Pfeifer, D., Zhang, Y.S. (1989). A survey on strong approximation techniques in connection with records. In: *Extreme Value Theory.* Proceedings, Oberwolfach 1987. J. Hüsler, R.–D. Reiß, Eds., *Lecture Notes in Statistics* **51**, Springer, N.Y., 50 – 58.

Rényi (1962). Théorie des éléments saillants d'une suite d'observations. Colloquium on Combinatorial Methods in Probability Theory, Mathematisk Institut, Aarhus Universitet, Denmark, 104—115.

Resnick, S.I. (1987). *Extreme Values, Regular Variation, and Point Processes.* Springer, N.Y.

Yang, M.C.K. (1975). On the distribution of the inter-record times in an increasing population. *J. Appl. Prob.* **12**, 148—154.

Vervaat, W. (1973). Limit theorems for records from discrete distributions. *Stoch. Proc. Appl.* **1**, 317 – 334.